Abstract

Pathogen nonpoint source (NPS) pollution is a leading cause of surface water impairment in the United States. Often, the sources of NPS pollution are difficult to ascertain. Previous studies have employed land use regression methods to develop a greater understanding of the sources of microbial NPS pollution. The results indicate that urban development and percentage of impervious surface are correlated with higher fecal coliform concentrations. In addition to the land cover and urban development variables, this study seeks to find models of fecal coliform contamination based on the spatial coverage of federal and state stormwater management policies. Such policies have not previously been included in predictive models, even though they could have an impact on water quality. Furthermore, the model we construct is on a larger, state-level geographical scale that spans multiple watersheds as opposed to the single watershed models commonly found in the literature. Nine parameters from over 30 predictor variables were selected for the final model, which has an adjusted R^2 of 0.61. This model can be used to improve predictions of fecal coliform levels at unmonitored freshwater locations across North Carolina. The model also presents the first steps in examining the large-scale effects of stormwater regulations on surface water pathogen levels, though more research is needed.

Table of Contents

Figures and Tables

1. Introduction

Initially, when the National Pollution Discharge Elimination System (NPDES) was created in 1972 as part of Section 402 of the Clean Water Act (CWA) (P.L. 92-500), it focused on point sources of surface water pollution. Since the location of point sources of pollution are known, their monitoring and control is straightforward. On the other hand, an area of growing concern is the effect of nonpoint source (NPS) pollution on water quality. NPS pollution is composed of contaminants washed off the surface of the land by stormwater runoff. Since NPS pollutants are diffuse in nature and do not enter water bodies from a single known point, identifying and mitigating the source poses many challenges.

The difficulty in controlling NPS contaminants is especially concerning, since NPS pollution is a leading cause of impaired surface waters in the U.S. (Arnold and Gibbons, 1996). The 2004 National Water Quality Inventory: Report to Congress (EPA, 2009) states that pathogens (bacteria) are the number one cause of impairment for rivers/streams and estuaries. This continued to be the trend in 2010, with The National Water Quality Assessment Report (EPA, 2010) stating that pathogens are the leading cause of impairment for impaired surface waters. The report for North Carolina states that fecal coliform pathogens are the fourth cause of impairment for rivers and streams, contributing to 10.9% of impaired water miles.

Pathogen contamination is measured by fecal indicator bacteria, which are bacteria present in the feces or other wastes of humans and other warm-blooded animals. Sources of fecal contamination to surface waters include wastewater treatment plants, on-site septic systems, domestic and wild animal feces, and municipal sewer overflows after heavy storm events (EPA, 2012). Fecal coliform bacteria are not usually harmful; however, they are used as indicators of bacteria which can be harmful to human health. Exposure to pathogenic bacteria from recreational contact with natural water can result in the following health effects: skin rashes, eye and nose infections and acute gastrointestinal illness, including cramps and diarrhea (Pruss, 1998).

Fecal indicator bacteria are used, because it is difficult, time-consuming, and expensive to test for a large variety of specific pathogens (EPA, 2012). Recently, the EPA has started recommending that states monitor surface water pathogens using *Escherichia coli* or enterococci instead of fecal coliform, as they are more highly associated with human health risk (EPA, 2012). However, since fecal coliforms have been the standard required for monitoring and reporting microbial contamination to the EPA in the past, data for those bacteria are most prevalent across space and time and thus were selected for use in this study.

The lack of knowledge on the source and transport of NPS contaminants, including fecal coliforms, has led to a growing body of scientific research aimed at predicting water quality degradation from land cover variables, including level of development, urbanization, and impervious surface coverage (Mallin et al. 2009). Most previous studies of NPS pollution have been focused on "low order streams in small catchments with a wide variety of land uses sampled over short time periods (<1 year)," (Baker, 2005, p. 2). Similar studies have not been conducted at a larger multi-watershed, state-level scale.

One of the main contributions of the research presented in this technical report is the investigation of the relationship between land cover and fecal coliform contamination at a large, state-level geographical scale. North Carolina has surface water monitoring locations across the state. The spatial and temporal extent of surface water monitoring in North

Carolina allows for the use of geospatial statistical methods to predict water quality at unmonitored locations. Thus, this study employs a land use regression analysis across the entire state. The analysis is performed in order to determine if fecal coliform contamination is related to land development characteristics such as land cover, impervious cover, road network density, population and housing density and different stormwater management policies at the state level.

Another significant contribution of this research is the incorporation of state and federal stormwater policies into a predictive regression model. No previous studies were found, which incorporate stormwater management policies into a predictive model of surface water quality. However, there is a good chance that such programs could have an impact on contaminant levels. This study specifically includes policy variables in the model selection process and found that inclusion of policy variables helped improve model fit by about 5%.

The final multivariate model has the potential to improve predictions of fecal coliform at unmonitored locations over simple interpolation methods. Additionally, the model provides useful information to land-use planners, resource managers and regulatory agencies regarding the current conditions of water quality across the state of North Carolina. This information can help in prioritizing future mitigation and restoration efforts.

For example, we used the predictive model derived in this study to analyze different local government stormwater best management practices (BMP) implementation methods across the state. The findings indicate that charging fees for noncompliance in long-term BMP maintenance responsibilities is associated with lower fecal coliform levels. However, more research in this area is needed.

2. Background

2.1 Land Use and Fecal Coliform

Many studies have investigated the link between land cover or impervious surface and water quality. Barbec et al. (2002) provide a review of articles which link level of impervious surfaces to water quality. Arnold and Gibbons (1996) also look at the relationship between impervious cover and water quality and provide strategies to reduce impervious surfaces for community planning, site-level planning and design, and land use regulation. Baker (2005) synthesizes literature investigating the relationship between land use and water quality, both for urban and agricultural impacts, looking at nutrient, chemical and bacterial contaminants. One conclusion he makes is that, "in general, as soon as urban land use increases beyond a small percentage of total land cover (~5%; the precise amount requires further research), then urban land use impacts on water quality dominate over agricultural" (p. 5).

Other studies have looked at measures such as population density (Peierls et al., 1991; Frenzel and Couvillion, 2002), housing density (Young and Thackston, 1999; Goonetilleke et al. 2005; Lohse and Merenlender, 2009), septic tank density (Kelsey et al., 2004; Cahoon et al., 2006), and commercial, industrial, and residential zoning (Selvakumar and Borst 2006; Mallin et al. 2009, Harclerode et al., 2012). Studies have also looked at changing land use and water quality within a watershed over a long period of time (Ren et al. 2003, Mehaffey et al. 2005). Figure 1 shows the relationship between impervious surface and stream health based on the research by Schueler (1992), which was updated in 2009 (see Shueler et al. 2009).

Figure 1. Stylized relationship of imperviousness to stream health (modified from Schueler 1992), from Arnold and Gibbons (1996), p. 246.

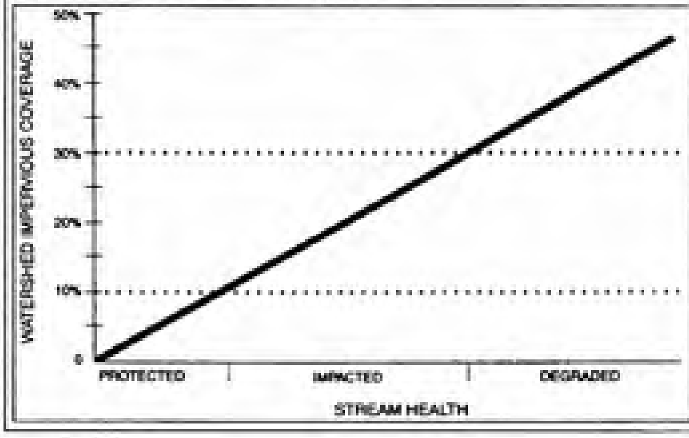
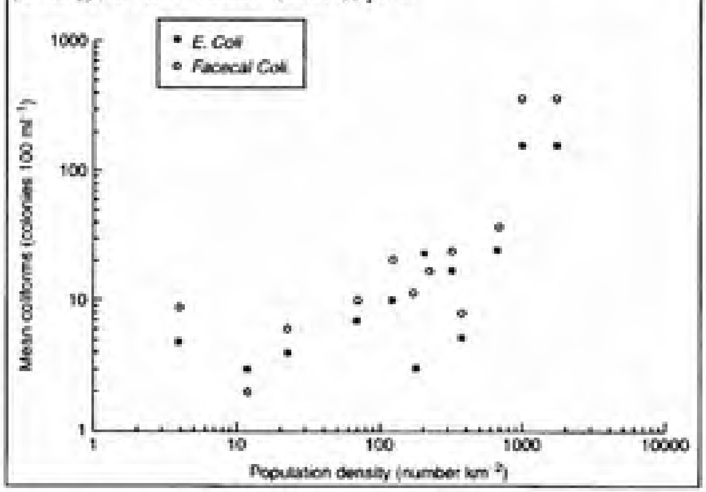


Figure 2. Graph of mean E. Coli and Fecal Coliform against population density (After Frenzel and Couvillion (2002)); From: Baker (2005), p. 3.



Specific to fecal coliform contamination, Frenzel and Couvillion (2002) showed that there is a correlation between fecal bacteria and population density as shown in Figure 2. Another study by Sliva and Williams (2001) showed correlations between urban land use and fecal coliforms within urbanized catchments. Young and Thackston (1999) found that tributaries in sewered basins had higher bacteria counts than those in non-sewered basins and that "fecal bacteria densities were related to the density of housing, population, development, percent impervious area, and apparent domestic animal density" (p. 1177). Mallin et al. (2000) conducted a study of North Carolina coastal water quality, finding that fecal coliform abundance was significantly correlated with the population and the percentage of developed land within the watershed, and most strongly with the percentage watershed-impervious surface coverage (p. 1047).

In a more recent study, Mallin et al. (2009) also studied the effects of different land uses to bacterial water quality. In addition to the typical land use categories of "developed land, percent impervious surface coverage, percent agricultural (including grazing land) and forestry combined," they also included a commercial and industrial use category which combined "percent of land devoted to retail business, services, manufacturing, and institutions (such as university facilities)" (p. 480). The results show that non-residential development was correlated with fecal coliform concentrations. However, they also found that impervious surface and overall urban development were not significantly correlated to fecal coliform concentrations, which goes against similar studies. The authors suggest this contradiction is likely due to the analytical limitations of their study.

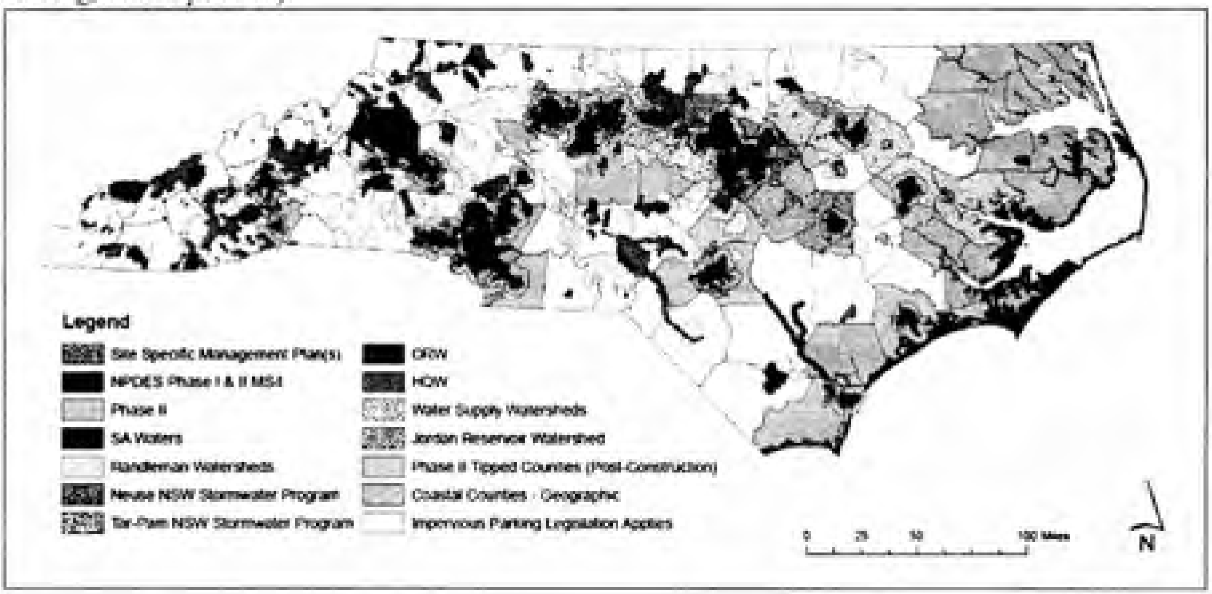
As for land covers found to protect water quality, Tong and Chen (2002) found that fecal coliform bacteria were negatively related to forest land use. However, a study by Smith et al. (2001) revealed a positive relationship between forest and wetlands along streams with fecal coliform. This finding was consistent with Fisher et al. (2000), who also found that there were high levels of fecal coliform in some predominately wooded watersheds. Both studies hypothesize that these relationships were a result of wastes from wildlife populations.

2.2 Stormwater Policies and Fecal Coliform

Since NPS pollution was declared as the nation's leading threat to water quality, programs to control surface water runoff have been devised by federal, state and local governments. For example, North Carolina has many state laws which regulate stormwater for new development, as seen in Figure 3, and as listed here: (1) the Neuse River Basin and Tar-Pamlico Basin Nutrient Sensitive Waters Management Program (requires nutrient-reducing BMPs); (2) Coastal Counties (requires Coastal Areas Management Act (CAMA) permit or a Sedimentation/ Erosion Control Plan, and special measures for shellfish areas); (3) High Quality Waters and Outstanding Resource Waters (requires a Sedimentation/

Erosion Control Plan); (4) Water Supply Watershed Protection Program (BMP and maximum density requirements based on water supply classification); (5) 401 Water Quality Certification (applies to wetlands an water of the State); (6) Universal Stormwater Management Program (voluntary program); (7) Impervious Vehicular Parking Legislation (applies to all areas which do not fall under any other state or federal regulations), and (8) National Pollutant Discharge Elimination System NPDES Stormwater Phase I and Phase II.

Figure 3. *All North Carolina Stormwater Permitting Programs as of June, 2009 (from NCDWQ, 2009, p. 2-13)*



Perhaps the best known NPS pollution policies are the federal NPDES stormwater program's Phase I and Phase II components. Phase I was implemented in 1990 for all municipalities with a population of over 100,000 and which operate a Municipal Separate Storm Sewer System (MS4). The focus of the regulations was onsite and operations planning to reduce nonpoint pollutant sources. Phase II was established in 1999, but didn't get instigated in NC until 2005, for smaller communities that own and operate a MS4.

Stormwater regulations mandate that many North Carolina local governments comply with the following six minimum measures of the Phase I and II regulations (NCDWQ, 2009):

1. Public education and outreach on stormwater impacts
2. Public involvement/participation
3. Illicit discharge detection and elimination
4. Construction site stormwater runoff control
5. Post-construction stormwater management in new development and redevelopment
6. Pollution prevention/good housekeeping for municipal operations

While the extent to which all six measures are met is important to reducing NPS pollution and improving water quality, this project specifically looks at the fifth measure: the implementation of post-construction stormwater management in new development and redevelopment. It is this measure that mandates communities to adopt ordinances that may require the use of structural best management practices (BMPs) for stormwater control. Specifically, post-construction stormwater management applies to developments (or redevelopments) in which the total land disturbance is one acre or more. The NPDES stormwater program classifies development into categories of low-density and high-density, both of which require a permit. (NCDWQ, 2009)

There are many complexities involved with the implementation, monitoring, maintenance and enforcement of structural stormwater management programs. According to Ellis and Marsalek (1996), "The success and sustainability of BMPs has to be ensured through proper design, operation, and maintenance to meet specific objectives" (Barbec et al., 2002, p. 510). A Survey of Local Government Post-Construction BMP Maintenance and

Enforcement in North Carolina (Barnes and Bruce, 2008) (henceforth "BMP Survey"), found that "local governments vary significantly in the ways in which they oversee the planning, installation, and monitoring of BMPs" (p. 3). This variability in stormwater management provides a basis for comparison to assess which practices by local governments are more successful in reducing fecal coliform NPS pollution and improving water quality.

The effectiveness of stormwater runoff controls in North Carolina has been assessed for agricultural BMPs. Line (2002) conducted an eight-year study of surface water in the Long Creek watershed in North Carolina. The watershed had undergone considerable land use and management changes during the study period, including implementation of agricultural BMPs. The study provided "evidence to suggest that the implementation of BMPs in the watershed have significantly reduced phosphorus and bacteria levels in Long Creek" (p. 1691). Santhi et al. 2006 analyzed the long-term effects of water quality management plans and BMP implementation on reducing nonpoint source pollution in an agricultural watershed in Texas. This study used simulations of contaminants rather than actual field measurements. Few studies have attempted to evaluate the effectiveness of BMPs to reduce NPS pollution from urban development at a large watershed scale.

3. Data

3.1 Fecal Coliform Data

The source of fecal coliform data for this study was the EPA STORET Database. This database stores water quality monitoring data collected by a variety of water resource management groups across the country, including states, tribes, watershed groups, other federal agencies, volunteer groups and universities. Data in STORET must contain a certain level of metadata, including where, how, why, when and what was monitored, to ensure data quality. Other studies which have used STORET data include Fisher et al. (2000), Wang (2001), and Tong and Chen (2002).

Data were queried for the state of North Carolina, from January, 1, 2002 to December, 31, 2010 for all surface water monitoring locations, including: streams, rivers, canals, lakes, reservoirs, wetlands and estuaries. Measurements were in colony forming units (cfu) per 100 milliliters. If cfu counts were reported as being below the detection limit, a value of 0.25 cfu/100ml was assigned. This value was taken as the midpoint between 0 and the lowest reported value of 0.5 cfu/100ml.

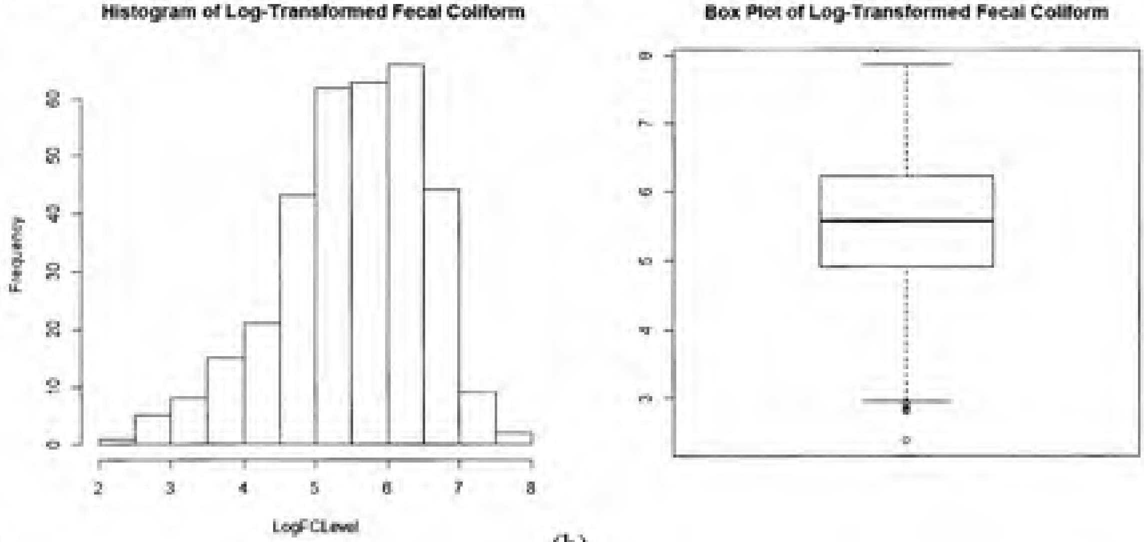
Fecal coliform distributions can be highly variable in the stream environment. According to the EPA, sampling bacteria from natural waters poses challenges, since "natural bacteria levels in streams can vary significantly; bacteria conditions are strongly correlated with rainfall, and thus comparing wet and dry weather bacteria data can be a problem; [and] many analytical methods have a low level of precision yet can be quite complex" (EPA, 2012). Additionally, following a routine sampling regime, as was the case for many of the monitoring locations, "allows for the equal chance of capturing a storm event or a non-storm event" (Schoonover and Lockaby, 2006). This issue is important since fecal coliform levels can dramatically increases after storm events. To account for sampling and seasonal variability, only monitoring stations that had at least 9 samples per year were included in the analysis. There was an average of 111 samples per location and a median of 105 [range: 94 - 217]. The samples at each location were averaged over the nine-year period. As is common practice, the means were log-transformed to reduce skewness and provide a more normal distribution of the data as seen in Figure 4a (see Silva and Williams (2001), and Mehaffey et al. (2005) of other studies which also applied a log transformation). The data were also checked for outliers as shown in the box-plot in Figure 4b. While there were a few suspicious observations, closer inspection showed that the fecal coliform levels were plausible and the points were not influential in altering the regression analysis. Thus, these points were not excluded from the analysis.

Summary statistics for the 418 selected sites can be seen in Table 1. The spatial location of each site is shown in Figure 5. Since previous studies, such as Mallin et al. (2000), found that there was a strong inverse relationship between salinity and fecal coliform levels, the data were split into freshwater and estuary categories. This was done because the majority of the observations lacked reliable salinity data. As Table 1 shows, the mean fecal coliform level for freshwater sites was much higher than the mean for estuaries.

Table 1. *Summary Statistics for fecal coliform observations (cfu/100ml) at 418 monitoring stations across North Carolina from 2002-2010.*

	N	Mean	St. Dev.	Median	Min	Max
Total Observations	56,085	303	1132	55	0.25	60000
Station Averages	418	316	329	210	0.87	2597
Fresh Water	339	368	331	265	10.8	2597
Estuaries	79	90	202	36	0.87	1627

Figure 4. Data Distribution: (a) Histogram and (b) Boxplot of Log-transformed Average Fecal Coliform at 339 Freshwater Monitoring Stations



(a) (b)

Figure 5. Selected Surface Water Sampling Sites in North Carolina (2002-2010)

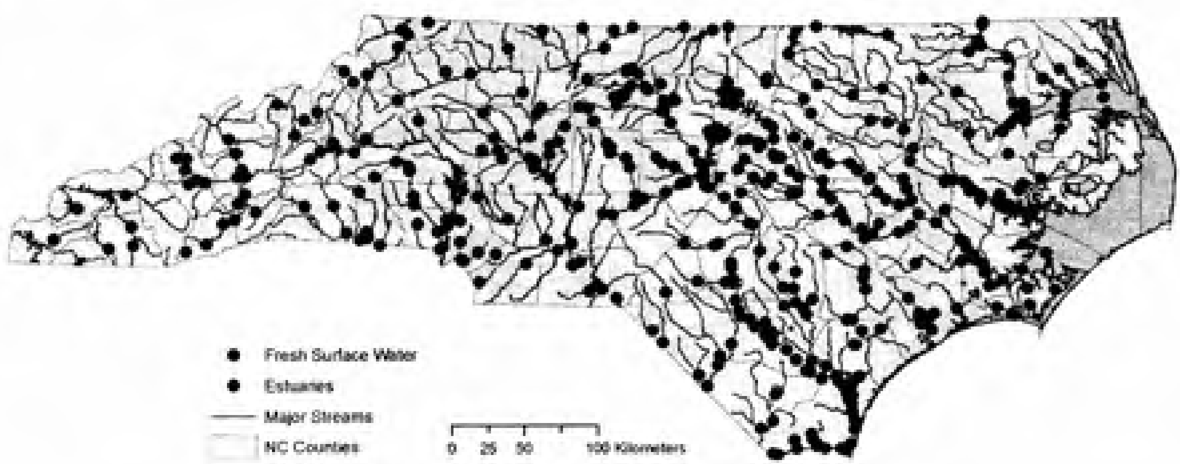
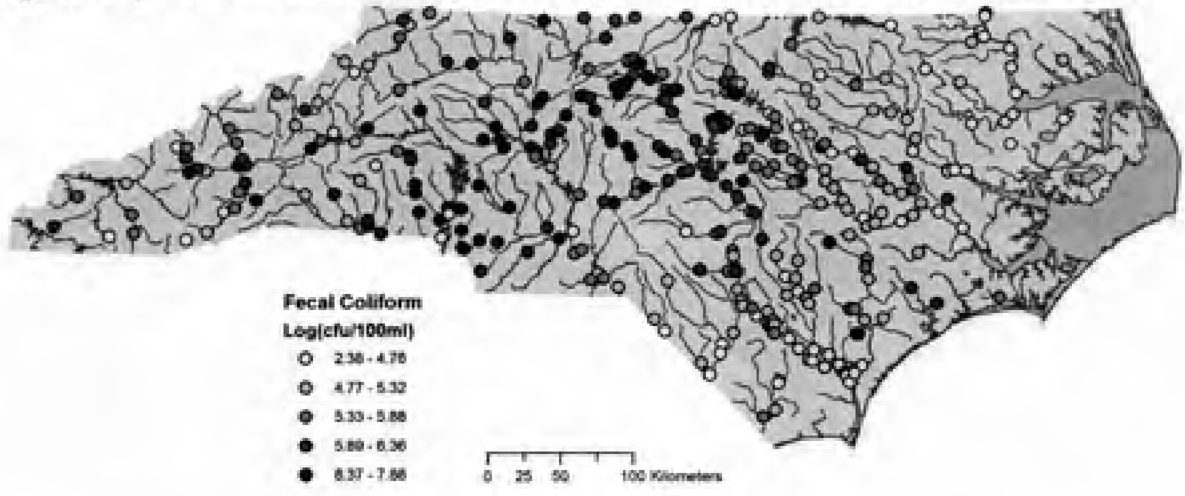


Figure 6 shows the spatial distribution of fecal coliform in fresh surface water across North Carolina for the study time period. It is evident that there is a strong geographical pattern in the data. Observations in the eastern part of the state tend to be lower than the rest of the state. Additionally, the central, most urbanized part of the state has the highest fecal coliform levels.

Figure 6. Spatial Distribution of Log(Fecal Coiform) Across North Carolina (2002-2010)



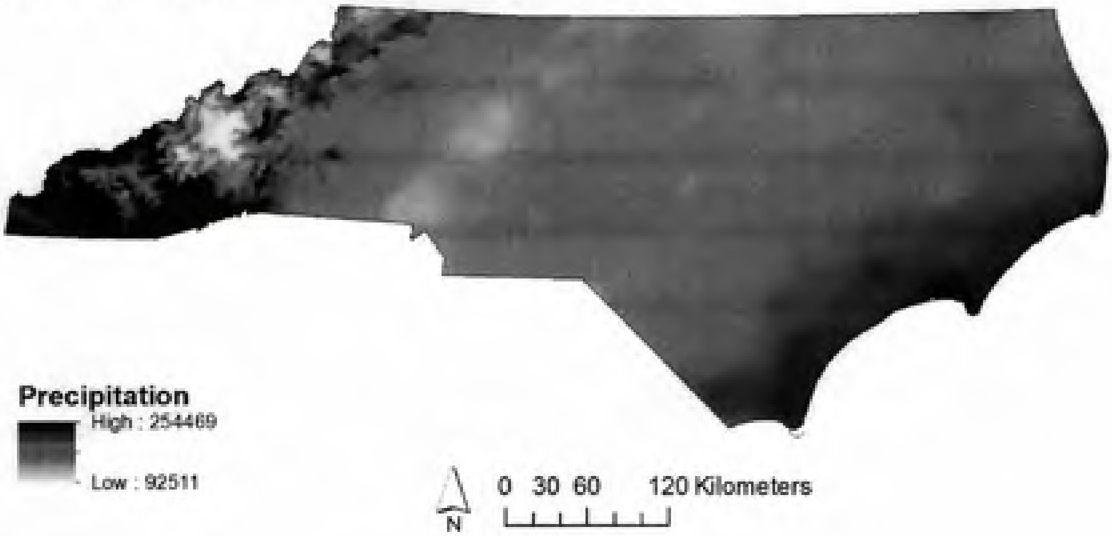
3.2 Explanatory Variables

3.2.1 Environmental Variables: Precipitation and Temperature

Mallin et al. (2000) noted that fecal coliform densities were strongly correlated with turbidity (positively) and salinity (negatively) and were subject to seasonal variations. Additionally, Kelsey et al. (2004) modeled the effects of land-use parameters by accounting for the effects of environmental variables such as tide, salinity, and temperature. Salinity was controlled for by excluding estuaries from the data set. Temperature data for the monitoring locations was queried from the EPA STORET database for the same time period as the fecal coliform samples. A map of temperature is not provided, but water temperatures are lowest in the western mountainous portions of the state and tend to be highest near the southeastern portions of the state.

It is a widely recognized that level of fecal coliform found in surface water of a nearby receiving stream increases after heavy precipitation (for example see Mallin et al. 2001). Average annual precipitation data for the time period of 1981-2010 was acquired from the PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate mapping system (PRISM Climate Group). PRISM uses point measurements of climatic factors to produce continuous, digital grid estimates. The precipitation normal data were available at an 800 meter grid cell resolution and are depicted in Figure 7.

Figure 7. *North Carolina Precipitation Normal for 1981-2010*

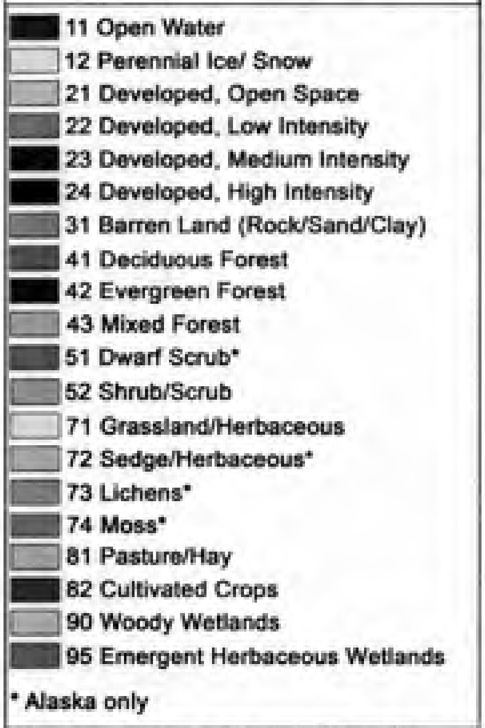


3.2.2 *Land Cover and Impervious Surface*

As mentioned previously, changes in land cover lead to changes in water quality (Gilliom et al., 2006, Wickham et al., 2008) and watershed runoff (Ponce and Hawkins, 1996). Thus land cover and impervious surface are key datasets for this model.

The National Land Cover Database (NLCD) holds data that show both natural and manmade land cover of the United States. These data are collected from Landsat satellites by the U.S. Geological Survey. The data used in this analysis is from the NLCD2006. This data set is a 16-class land cover classification scheme at a spatial resolution of 30 meters. The dataset also includes a layer for percent developed impervious surface, which is shown in Figure 10 (Fry, 2011).

Figure 8. NLCD Land Cover Classification Legend



Source: http://www.mrlc.gov/nlcd06_leg.php

The land classification is based on the Anderson Land Cover Classification System. The 16 classes for the conterminous United States are summarized in Figure 8 and mapped for North Carolina on Figure 9. When developing the statistical model for fecal coliform prediction, individual subclasses and grouped land cover classes were used. Aggregation of the subclasses into the grouped class was based on the first digit of the class code. For example, classes 11 and 12 were grouped together; classes 21, 22, 23, and 24 were grouped together; and so on. Grouped raster data files were created using the reclassify tool in ArcGIS10. Many studies which have analyzed the impact of land use on surface water quality have aggregated land cover in this or similar ways (e.g. Bolstad and Swank 1997, Smith et al. 2001, Kang et al. 2010, Mallin et al. 2009, Pratt and Chang, 2012).

Figure 9. *Land Cover in North Carolina (NLCD 2006)*

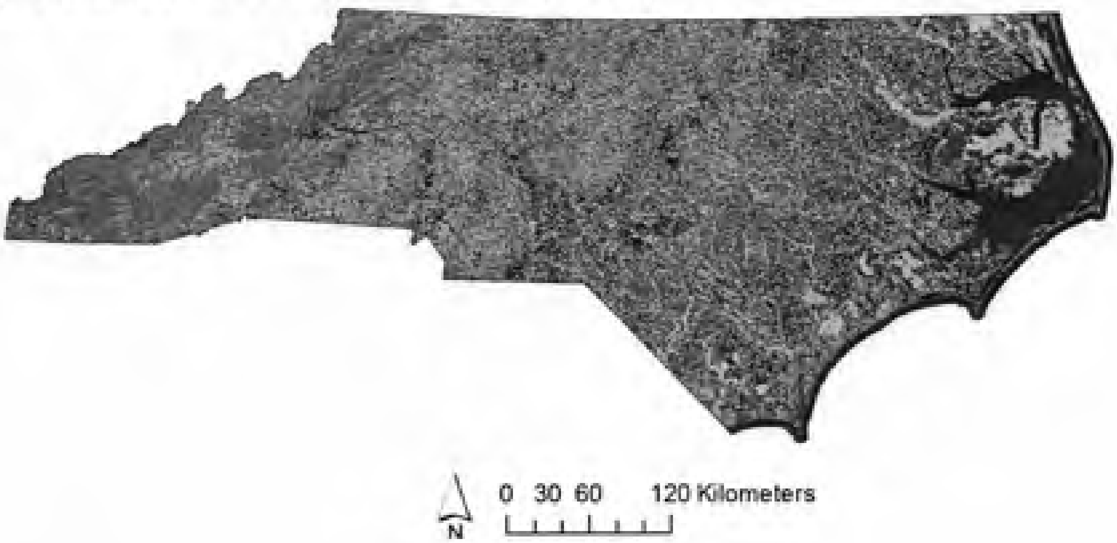
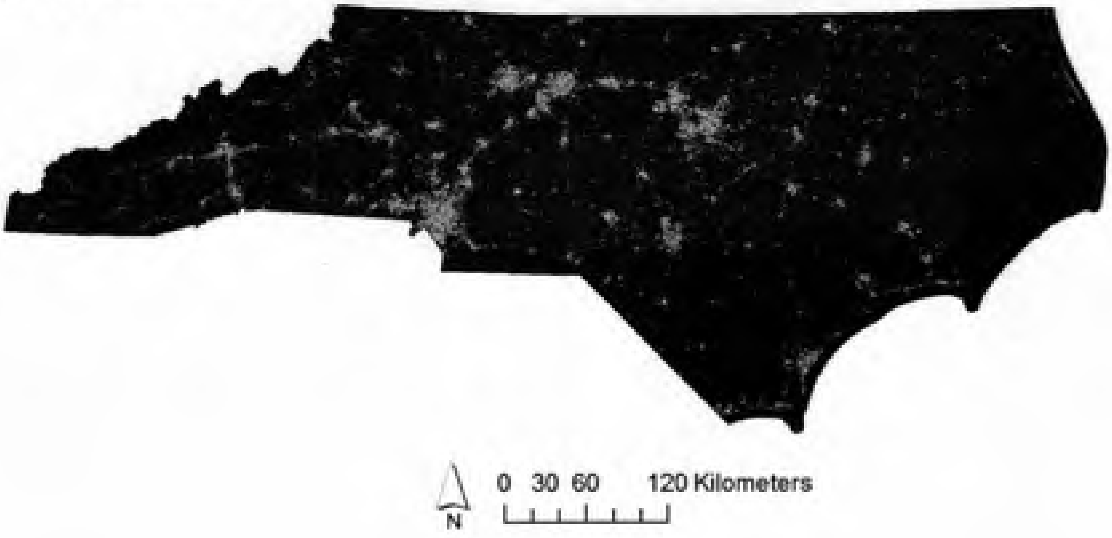


Figure 10. Percent Impervious Surface in North Carolina (NLCD 2006)



3.2.3 Housing, Population and Road Density

In addition to the "developed" land cover classes, housing, population and road network data were used to represent the level of urban development in a geographical area. Housing and population data were retrieved from the 2010 Census TIGER/Line Shapefiles at the census block level. Density for each variable was calculated by dividing the counts by the area of the census blocks. Figure 11 shows the population density. A map of housing density is not shown, since it is nearly identical to the population density map.

Integrated statewide road network data were from 2007 and from the North Carolina Department of Transportation (NCDOT). The road network density was calculated at the scale of the 2010 census blocks used in the housing and population density calculations, and can be seen in Figure 12. Comparing Figure 11 and Figure 12, it can be seen that these variables are highly correlated with each other. The high multicollinearity exhibited in the different variables is further evaluated in the multivariate regression model.

Figure 11. Population Density in North Carolina (2010 Decennial Census)

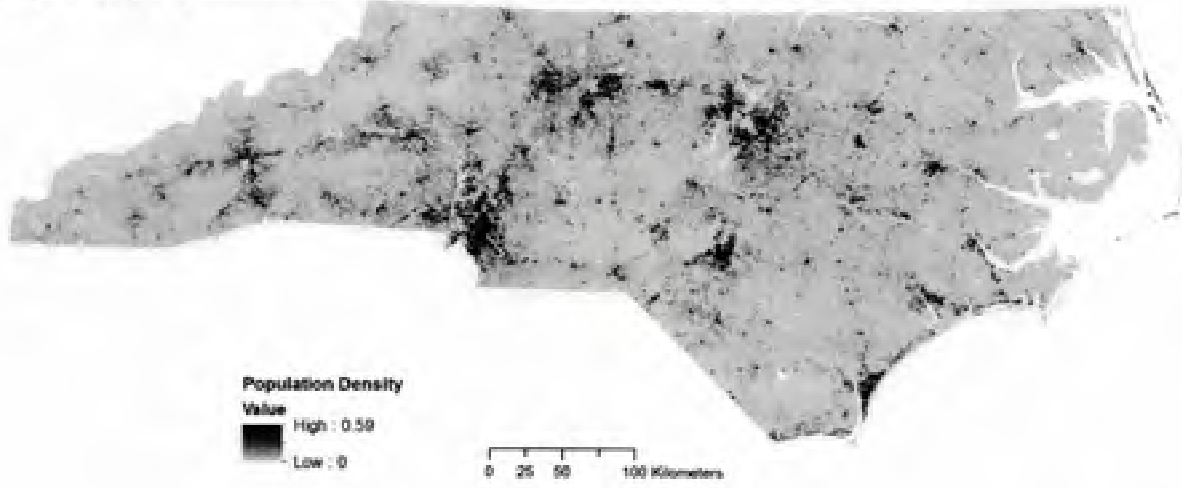
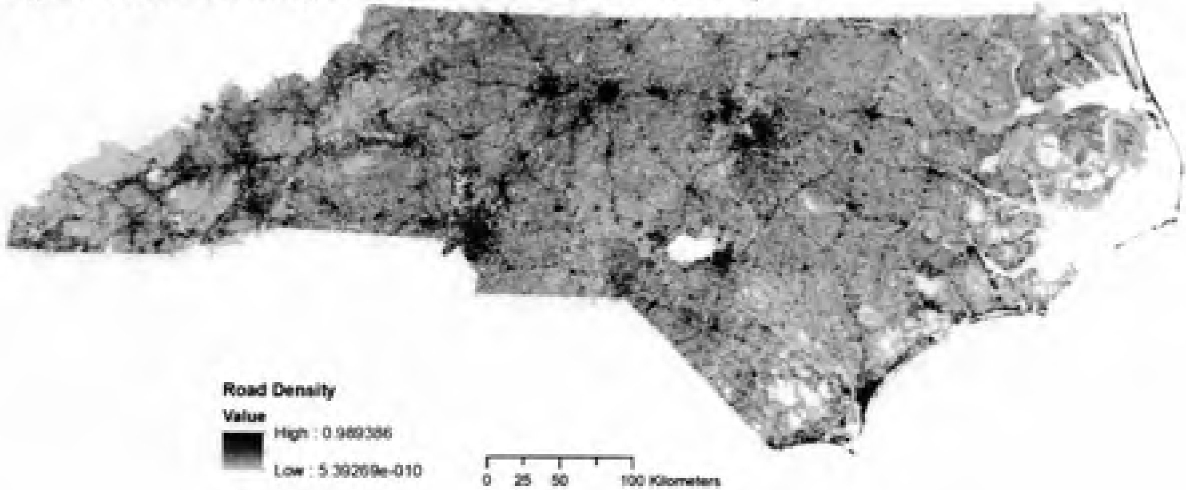


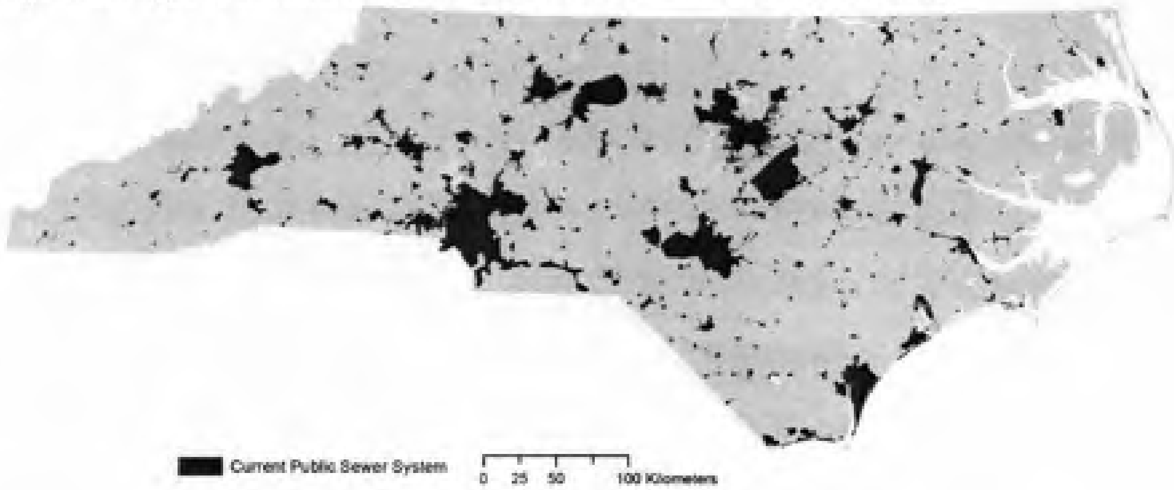
Figure 12. Road Density in North Carolina (NCDOT 2007)



3.2.4 Sewer District Coverage

Some previous studies which have modeled fecal coliform contamination have looked at septic tank density (Kelsey et al., 2004; Cahoon et al., 2006). This data could not be located at the statewide scale for North Carolina. Instead, an alternative measure was used: population within Type A Current Public Sewer Systems (CPSS). Type "A" refers to sewer systems which serve the general public, accept domestic wastewater and are generally considered large systems. This data was retrieved from the NC OneMap Geospatial Portal. It was created by The NC Center for Geographic Information and Analysis and mapped by contractors to the NC Rural Center (engineering firms McGill & Associates and Hobbs, Upchurch & Associates) during 2004, 2005, and 2006. The data is complete for all of the 100 individual counties of North Carolina. "Current" refers to the most recent year of data the sewer system owner had that represented a full year. Since the survey was in 2004, this data would most likely have been for calendar year 2003. The CPSS in North Carolina as of 2003 are shown in Figure 13.

Figure 13. *Type A Current Public Sewer Systems in North Carolina (2003)*

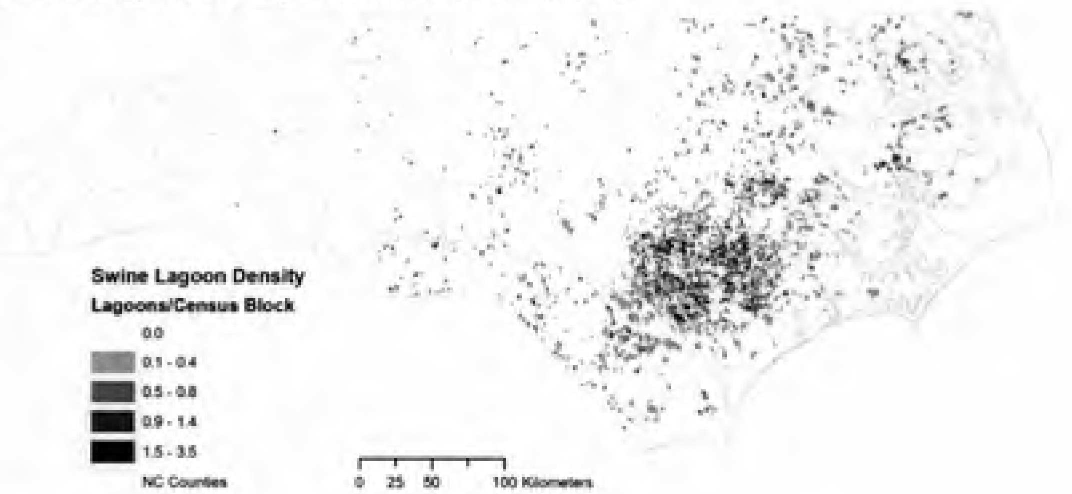


<u>3.2.5 Swine Lagoon Density</u>

Hooda et al. (2000) list farm effluents (livestock wastes) and bacterial contamination of water as main concerns regarding water quality degradation related to livestock farming areas. According to the 2007 USDA Agricultural Census, North Carolina ranks as the second-largest hog farming state in the country, after Iowa (USDA, 2007). Thus, an important variable to consider is the density of swine lagoons. Point locations of intensive animal operations in North Carolina were collected in 1998 by the North Carolina

Department of Environment and Natural Resources, Division of Water Quality. The data were amended in 2003, by CGIA through use of digital ortho photography from 1998 to verify the DWQ points for swine operations and to identify additional swine lagoons and barns visible in the photos. The point data was used to calculate swine lagoon density at the 2010 census block scale. The results are shown in Figure 14.

Figure 14. *Swine Lagoon Density in North Carolina (1998)*



3.2.6 Statewide Stormwater Policies

There are many stormwater control programs designated by federal, state and local governments. Figure 3 in the Background section shows the spatial distribution of the different stormwater permitting regulations across North Carolina as of June 2009. The programs were grouped in the following way and included as separate variables in the model: (1) Coastal Counties (CAMA); (2) High Quality Waters (HQW) and Outstanding Resource Waters (ORW) outside the 20 coastal counties; (3) Water Supply Watershed Protection Program; (4) Nutrient Management Programs (NSW); (5) NPDES Stormwater Program (Phases I and II); and (6) None.

3.3 2006 Post-Construction BMP Survey

In 2007, Triangle J Council of Governments and the UNC Environmental Finance Center conducted a survey of local North Carolina governments on financing, maintenance and enforcement of post-construction BMPs (Barnes and Bruce, 2008). The survey asked over 30 questions about local government BMP implementation practices, and received responses from 58 jurisdictions. Three responding jurisdictions could not be identified and thus were dropped from analysis. Additionally, six municipalities within the same county had identical responses for all but two questions, and thus all were excluded to avoid biasing the results. Responses from 14 counties and 35 municipalities were included in the analysis. The 49 jurisdictions were matched to county and city political boundaries by name. A spatial join between the political boundaries and monitoring stations was conducted in ArcGIS to select a subset of 75 points from the original 339 that fell within the surveyed jurisdictions. Due to confidentiality constraints of the survey, maps of this data are not presented in this report.

There are a few caveats of the survey data. First, the response rate for the survey was low at 36%. Also, the researchers focused on contacting larger more established stormwater programs to increase response. The authors state that "because of this fact and because larger systems are more likely to be able to devote time to surveys, larger jurisdictions may be more heavily represented in the survey results" (p. 6). Second, the authors also mention that survey respondents from small jurisdictions with very new stormwater management programs stated that their responses should not be viewed as representative of long-term policy, as their programs were still in the formative stages of development (p. 3).

The questions whose responses were selected to be included in the model selection are listed below. These questions were selected based on the response rate as well as how relevant their scope was to directly impacting water quality. For a full list of questions and details on response rates in the original survey, see Barnes and Bruce (2008).

Q4. Does your jurisdiction assess an ongoing stormwater fee on developed properties, based on amount of impervious surface or some other factor?
Q7. How does your program review/approve/require stormwater BMPs for new development?
Q9. Stormwater plan review in your jurisdiction is funded primarily from…? *[modified: fees, general fund, or a mix?]*
Q10. Do you field-verify BMPs on private land after construction in your jurisdiction? *[modified]*
Q12. Do you have a centralized inventory or jurisdictional BMP map to track for each BMP in your jurisdiction? *[modified]*
Q14. For developments that include stormwater BMPs, what information about the BMPs must developers submit *before* their grading and/or construction permits can be issued? *[modified: long-term BMP maintenance plan or only preliminary BMP designs?]*
Q15. Do you require as-built inspections/certifications of BMPs after they are installed?
Q16. When BMPs for new development are inspected upon completion of construction, approximately what percentage of them does not comply with regulations or specifications and require some additional work?
Q18. Who inspects stormwater BMPs over the long term?
Q26. Do you investigate and/or track whether a given dysfunctional BMP caused a visible or measurable impact to surface waters?
Q27. Does your program or ordinance require entities responsible for maintenance to post financial and/or performance guarantees (e.g., bonds, letters of credit) for maintenance of BMPs after construction is completed and certificates of occupancy are issued?
Q30. What penalties for noncompliance with long-term BMP maintenance responsibilities does your program or ordinance include?
Q31. If you can assess penalties, what is the maximum amount per day that can be assessed according to your local regulations? *[modified: less than $5,000 or more than or equal to $5,000?]*

4. Methodology

4.1 Land Use Regression Methods

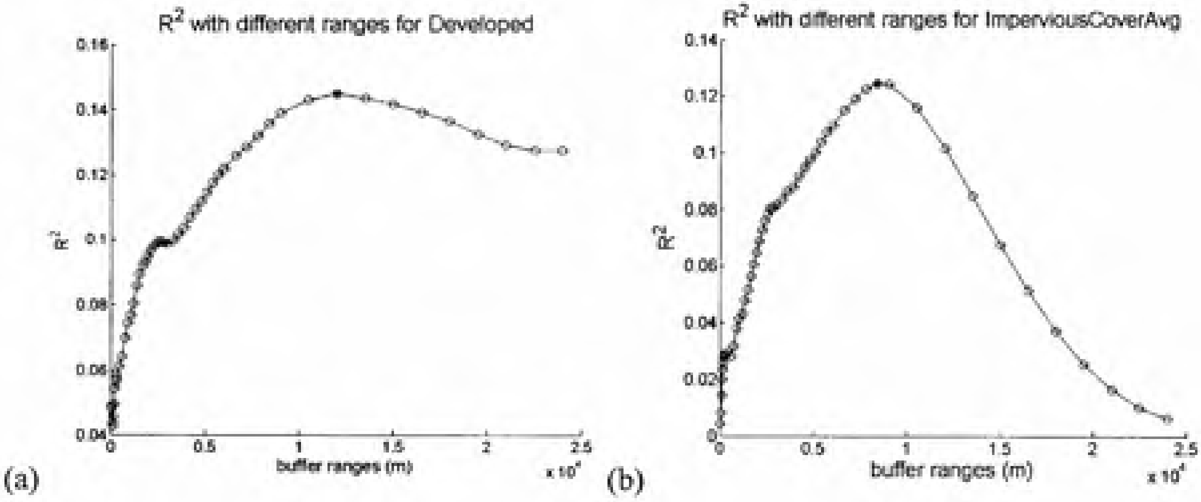
A land use regression method uses fixed-site water pollution monitoring data as a dependent variable and geographic information system (GIS) land use descriptor variables as independent variables to develop regression models, in order to predict fecal coliform levels at unmonitored locations and to elucidate relationships between contaminant levels and land variables. An increasing number of studies have employed this method for water quality (e.g. Wang 2001, Kelsey et al. 2004, Mehaffey et al. 2005, Kang et al. 2010, Yang and Jin 2010).

Two different methods were planned for this analysis: a buffer approach, and a subwatershed approach. The buffer method allows for easy exploration of relationships at different spatial scales, while the subwatershed level is more hydrologically accurate of pollutant transport. Silva and Williams (2001) conducted analysis of land use impact on river water quality at both the watershed scale and at a smaller 100 m buffer local scale. Their results indicate that both the smaller buffer and the watershed scale showed similar relationships with fecal coliform levels, though the watershed model performed slightly better. Bolstad and Swank (1997) and Pratt and Chang (2012) also analyzed land cover at both the catchment and smaller riparian buffer scales. Due to time constraints, the subwatershed analysis was not able to be included in this report. Instead, the buffer method is described below.

The first step of the buffer method was to create circular buffers around each monitoring location, which serve as the geographical unit of analysis for calculating the value of each explanatory variable (e.g. percent impervious surface). Buffer radii ranged from 30 meters to 24 km. The size of the buffer was constrained to be a multiple of the minimum land cover grid cell size of 30 m. The next step was to select which buffer size for each explanatory variable should be included in the multivariate regression. This was done through a series of bivariate regressions in MATLAB (The Mathworks Inc., Natick, MA, USA).

The explanatory variables at each buffer size were regressed against the log-fecal coliform data and the R^2's for each buffer radius were plotted. The radius which maximized the R^2 was selected. Figure 15 provides example plots of the R2's for two explanatory variables. Table 2 in the Results section summarizes the buffer radii selected for all of the explanatory variables.

Figure 15. *Example Buffer Radius Plots for (a) Developed Land Cover (b) Percent Impervious Cover. The red point is the buffer which maximized the R^2.*



4.2 Multivariate Model Selection

All explanatory variables were included in forward, backward, and exhaustive, or all possible subsets, selection procedures. These processes were conducted using the Leaps package for the R version 2.15.2 (The R Foundation for Statistical Computing). In all three procedures, all possible models for every model size were regressed using Ordinary Least Squares; the model with the best fit was selected for each model size (number of parameters). From this set of models, the best overall model was selected based on the lowest Bayesian Information Criterion (BIC). The BIC combines model fit and model complexity (number of parameters) into one measure. It includes a penalty for models with extra parameters, thus optimizes model selections by selecting the model with the best fit with the fewest number of parameters. BIC was selected over other similar criterions since it penalizes the most for having extra parameters. This was important for this model in particular, since it included a large number of independent variables and it was desired that the final model did not contain an excessive number of variables which would lead to an over-fitted model.

If collinearity exists between predictor variables the resulting regression coefficients may not be unique. We examined collinearity between all of the variables selected for the best model through the variance inflation factor (VIF). VIFs "report how much of a regressor's variability is explained by the rest of the regressors in the model due to correlation among those regressors" (Craney and Surles, 2002, p. 392). While there are no set criteria for deciding when a VIF is too large, commonly used cutoff values include VIF $\geq$ 5 or VIF $\geq$ 10 (Ibid). However, these cutoff values may be considered very lenient in regards to the correlation among the independent variables. For example, a VIF of 10 implies that 90% of the variabilility in the independent variable is explained by the remainder of the independent variables in the model. Kelsey et al. (2004) and Aufdenkampe et al. (2006) used a VIF cutoff of 10. For the analysis in this report, a more conservative VIF cutoff of 2.5 was selected, ensuring that no more than 60% of the variability of one variable is explained by the other independent variables in the model.

4.3 Addressing Autocorrelation

Surface water monitoring locations are part of a connected hydrological stream network. Thus, it is likely that points located along the same stream are not independent from one another. To address this concern, a weights matrix was created as described by LeSage (2004). The weights matrix specifies dependence between observations. It is defined as an n x n spatial weight matrix W with elements Wij>0 for observations j = 1. . . n within a 25 km linear distance upstream to observation i. Monitoring stations without any upstream neighbors were assigned a 0 in the weights matrix. Observations with one neighbor were assigned a 1, while observations with multiple neighbors were assigned a normalized value so that the row sum would equal one. This weights matrix was included in the spatial regression function as follows:

$$y = pWy + X\beta + \varepsilon$$

p = regression coefficient for upstream neighbors of a point
W = weights matrix of upstream neighbors
y = log-fecal coliform at upstream neighboring observations
X = log-fecal coliform at observation
β = regression coefficient for observations X
ε = error term

An improvement on this simplified weights matrix would be to weight the neighbors by stream distance as opposed to equally, however there was not sufficient time to conduct such an analysis for this project. Additionally, it would be beneficial to test whether the assumption of autocorrelation between sample sites exists. In a study by Cunningham et al. (2009), the authors tested for independence among sample sites, and found that adjacent sites were actually statistically independent (p. 274).

4.4 Testing Survey Policy Variables

Data used from the BMP Survey (Barnes and Bruce, 2008) covered 49 jurisdictions, within which were located 75 monitoring stations. Thus, these variables could not be tested during the model selection process which included the 339 statewide observations. Instead two different methods were proposed to test whether any of the local BMP implementation practices contributed to improving the predictability of the model selected based on the statewide variables.

The first method was to conduct a regression on the residuals from the model which was selected for based on all 339 observations. In effect, the first regression model can be seen as acting as a control for all the statewide land use and stormwater policy variables and what is left over in the residuals is what cannot be explained by those control variables. Then, the residuals of the 75 observations are regressed to see if any of the BMP survey data can explain any of the remaining variation.

However, the regression on the residuals technique, though commonly practiced, is criticized as it can produce biased parameter coefficients since it does not take into account the possible correlation between the variables regressed in the first model and those in the second model (Freckleton, 2002 and King, 1986). However, I argue that this technique is appropriate in this case, since the data used in the secondary regression is only available for a small subset of the original observed data. However, caution should be taken when interpreting the coefficients of the selected parameters.

Due to the limitations of the regression on the residuals method, a second method was used to provide a comparison. The second method was to conduct a new multivariate regression with all variables, statewide and from the BMP survey, on only the subset of 75 observations. Due to the large number of variables in both methods, only forward and backward stepwise regression was conducted. As with the original regression model, BIC values were used to select the best model.

While this second method produces unbiased coefficients, it also has drawbacks. The main disadvantage to this method is that 264 valid observations (78%), must be excluded from the analysis. This exclusion of data can also lead to bias in the model, as it is not randomized, but based on which jurisdictions were able to participate in the BMP survey. Ideally, the best solution to this issue is to conduct expectation maximization. However, time does not allow for such a procedure to be conducted for this paper.

5. Results

5.1 Buffer Sizes

The radii of buffers around monitoring locations were selected by conducting a bivariate linear regression between log-transformed nine-year mean fecal coliform at every freshwater monitoring station and each of the 33 explanatory variables. The selected radii are reported in Table 2 (next page), along with the R^2, p-value, and the sign of the coefficient.

All of the urban development variables had a positive correlation with fecal coliform, as was seen in previous studies. Interestingly, swine lagoon density was inversely related to fecal coliform levels. This could be because that variable is indicative of rural development or a regional pattern of lower fecal coliform levels in the eastern portion of the state, rather than showing the effect of animal waste.

Of the stormwater policy variables, (1) coastal, (2) high quality and outstanding resource waters, (3) nutrient sensitive waters, and (4) the absence of stormwater management policies, were all associated with lower fecal coliform levels. Water supply watersheds and Phase I and II policies were positively associated with fecal coliform. The result for Phase I and II policies makes sense, since that policy only applies to urbanized areas, which are also positively correlated with fecal coliform levels. This is also why the absence of stormwater management policies is associated with lower fecal coliform, since those are rural areas with little urban development and impervious cover.

As for land covers, all grouped land covers, including forest, were associated with increased fecal coliform levels. The only exception was wetlands, which were negatively associated with fecal coliform. Wetlands are abundant in the eastern part of the state and could be inversely related to fecal coliform bacteria due to their dilution and filtration properties.

Of all the individual land class variables, open water had the highest R^2 and on its own explained about 29% of the variability in fecal coliform levels. It is interesting to point out that within the forest category, deciduous and mixed forest were positively associated with fecal coliform, while coniferous forest was actually negatively associated with fecal coliform. Similarly, within the agricultural category, the pasture/hay land cover had a positive association, while the crop land had a negative association with fecal coliform. This could be because pasture/hay has wastes from grazing animals, while crop land does not likely have the presence of many farm animals. Another explanation is the geographical distribution of crop land, which is located predominantly in the eastern portion of the state, where fecal coliform levels are lower, while pasture/hay tends to cover the western half.

5.2 Multivariate Regression Results

Once buffer sizes for each explanatory variable were determined, a series of multivariate regressions were conducted. Regressions were run both with and without the policy variables. Inclusion of the policy variables resulted in an improved adjusted model R^2 of 0.05. Both individual land covers and aggregated land cover groups were included in the model selection process. This was done because individual variables within aggregated land cover groups sometimes showed opposite signs, as was the case with the forest and agriculture land covers.

The best models from a forward, backward, and all possible subsets regression were examined. While the model selected by the all possible subsets method produced the lowest overall BIC and highest R2, it had major multicollinearity issues, as multiple variables exhibited a VIF of >10. In contrast, the model selected through the backwards stepwise regression had the fewest parameters, a good R^2, and low multicollinearity. In this model, the variable with the highest VIF (wetlands) had a VIF <2.5, and all remaining variables had VIF values of <2.

Figure 16 shows the BICs for the models from the backwards stepwise regression. The x-axis lists the 33 explanatory variables, while the BIC's for each model are plotted on the y-axis. The model with the lowest BIC is considered the best. The parameters selected for this model are signified by grey squares in the topmost row of the figure. The statistics for this model with the lowest BIC are reported in Table 3. Of note, (1) pasture/hay, (2) open water, (3) wetlands, (4) high quality and outstanding resource waters, and (5) nutrient sensitive waters, were all selected by the backward, forward, and exhaustive selection methods. This could indicate that these five variables are reliable estimators of fecal coliform contamination.

Figure 16. *Plot of BICs for Models Selected using Backward Stepwise Regression Method. The top row indicates the model which best balances complexity and fit.*

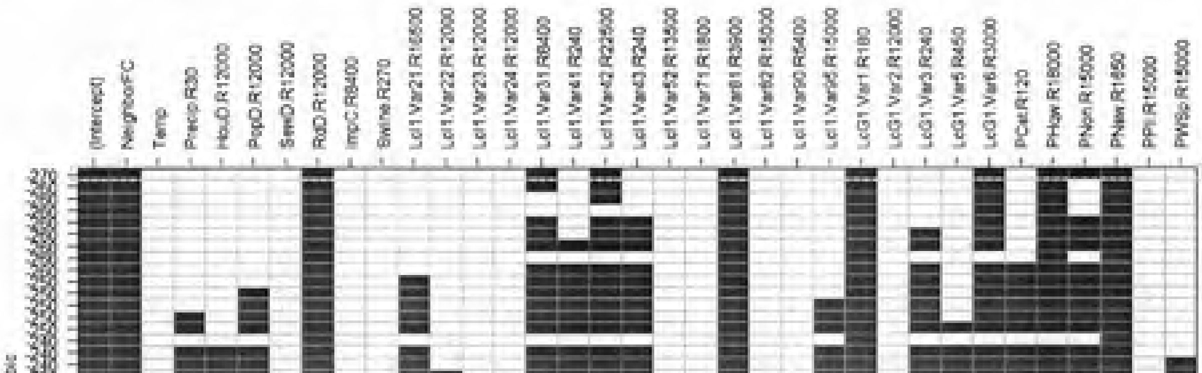


Table 2. Summary of Explanatory Variable Buffer Sizes

	Radius (m)	R^2	β sign	P-value
Urban Development Variables				
Housing Density	12000	0.1286	+	<0.01
Population Density	12000	0.1325	+	<0.01
Road Density	12000	0.2075	+	<0.01
Sewer Population Density	12000	0.1095	+	<0.01
Sewer System Coverage	12000	0.0884	+	<0.01
Impervious Cover	8400	0.1247	+	<0.01
Swine Lagoon	270	0.0215	–	0.01
Statewide Stormwater Management Policies				
	Radius (m)	R^2	*β sign*	*P-value*
Policy: Coastal	120	0.1603	–	<0.01
Policy: HQW ORW	18000	0.0333	–	<0.01
Policy: NSW (not sig)	1650	0.0078	–	0.10
Policy: Water Supply	15000	0.1110	+	<0.01
Policy: Phase I/II	15000	0.2230	+	<0.01
Policy: None	15000	0.0220	–	<0.01
Land Cover Variables				
	Radius (m)	R^2	*β sign*	*P-value*
Group: Developed	12000	0.1448	+	<0.01
Group: Forest	240	0.1292	+	<0.01
Group: Agriculture	450	0.0585	+	<0.01
Group: Wetlands	3000	0.3009	–	<0.01
LC11: Open Water	180	0.2949	–	<0.01
LC21: Developed Open	16500	0.1399	+	<0.01
LC22: Developed Low	12000	0.1448	+	<0.01
LC23: Developed Med	12000	0.1151	+	<0.01
LC24: Developed High	12000	0.1108	+	<0.01
LC31: Natural Barren	8400	0.0301	–	<0.01
LC41: Deciduous Forest	240	0.1563	+	<0.01
LC42: Coniferous Forest	22500	0.1341	–	<0.01
LC43: Mixed Forest	240	0.0223	+	0.01
LC52: Shrub	13500	0.1725	–	<0.01
LC71: Grassland	1800	0.0278	+	<0.01
LC81: Pasture/Hay	3900	0.2024	+	<0.01
LC82: Cultivated Crops	15000	0.1158	–	<0.01
LC90: Woody Wetlands	5400	0.2783	–	<0.01
LC95: Emergent Wetlands	15000	0.1983	–	<0.01
Environmental Variables				
	Radius (m)	R^2	*β sign*	*P-value*
Temperature	--	0.0333	–	<0.01
Precipitation	30	0.0836	–	<0.01

Table 3. Model Selected by Exhaustive Method for Aggregated Land Cover

	Buffer	Estimate	Std. Error	t value	
(Intercept)		5.3193	0.1846	28.82	***
Neighbor Fecal Coliform		0.098	0.0283	3.44	***
Road Density	(12 km)	194.675	34.3433	5.67	***
Natural Barren Cover	(8.4 km)	-28.081	10.1286	-2.77	**
Coniferous Forest	(22.5 km)	2.275	0.7482	3.04	**
Pasture/Hay	(3.9 km)	2.393	0.5154	4.64	***
Wetlands	(5.4 m)	-1.581	0.3818	-4.14	***
Open Water	(0.18 m)	-4.001	0.3343	-11.97	***
High Quality and Outstanding Resource Waters Policy	(18 km)	-2.411	0.4682	-5.15	***
Nutrient Sensitive Waters Policy	(1.65 km)	-0.659	0.1077	-6.12	***
No Stormwater Policies	(15 km)	-0.326	0.1185	-2.75	**

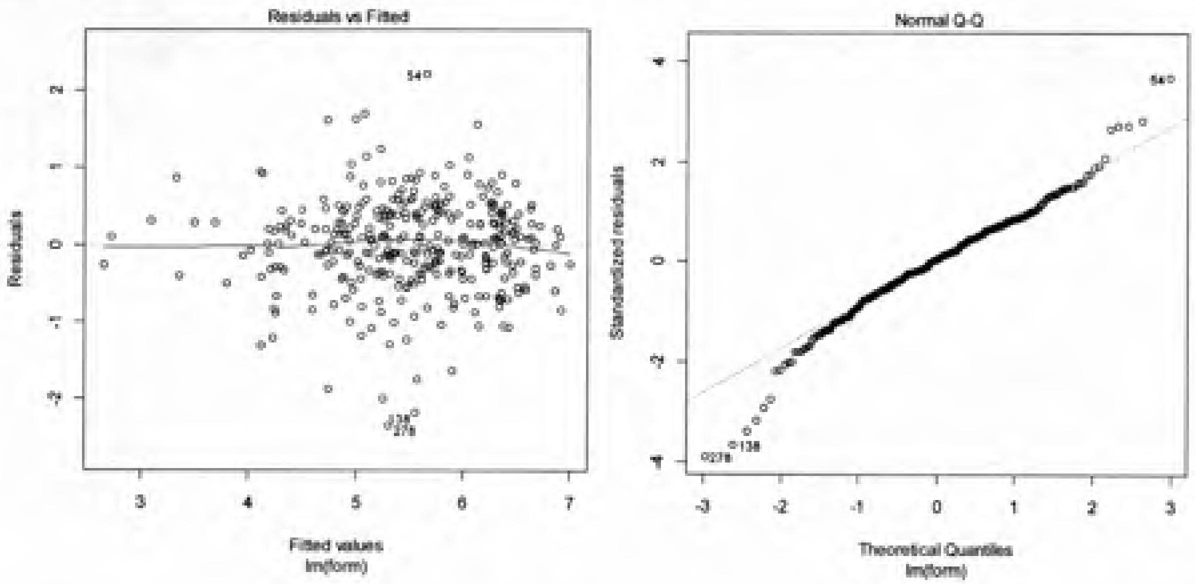
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Adjusted R-squared: 0.612, p-value: < 0.01

The multivariate model described in Table 3 has a total of nine explanatory variables. Of all the development variables, road density was the one selected for this model and showed a positive association with fecal coliform levels. Three policy variables were also selected for the multivariate model, all of which had a negative correlation with fecal coliform. Finally, five land cover variables were selected. Natural barren land, wetlands, and open water were all negatively associated with fecal coliform levels, while coniferous forest and pasture/hay were positively associated with fecal coliform. It is interesting to note that the coefficient for coniferous forest changed signs when the variable was included in the multivariate regression. A bivariate regression showed that coniferous forest was negatively associated with fecal coliform, but in the multivariate model a positive association was indicated.

5.3 Multivariate Regression Model Diagnostics

To ensure confidence in the model coefficients and accuracy of predictions, model diagnostics were run to determine whether the assumptions of linear regression were met. Plots of the residuals indicated that they exhibited a constant mean of 0 and followed a relatively normal distribution, as shown in Figure 17. Additionally, tests were run to identify if any points were highly influential on the model. One observation was identified as being influential, since it exhibited a Leverage value close to 0.5 and a Cook's Distance of about 0.5. When this observation was removed, the model R^2 increased and the residual standard error decreased. Furthermore, the original multivariate model included the variable for swine lagoons. However, after removing the influential observation, that variable was no longer selected as significant. Instead, the variable for no stormwater policies was selected. All other variables in the model remained unchanged.

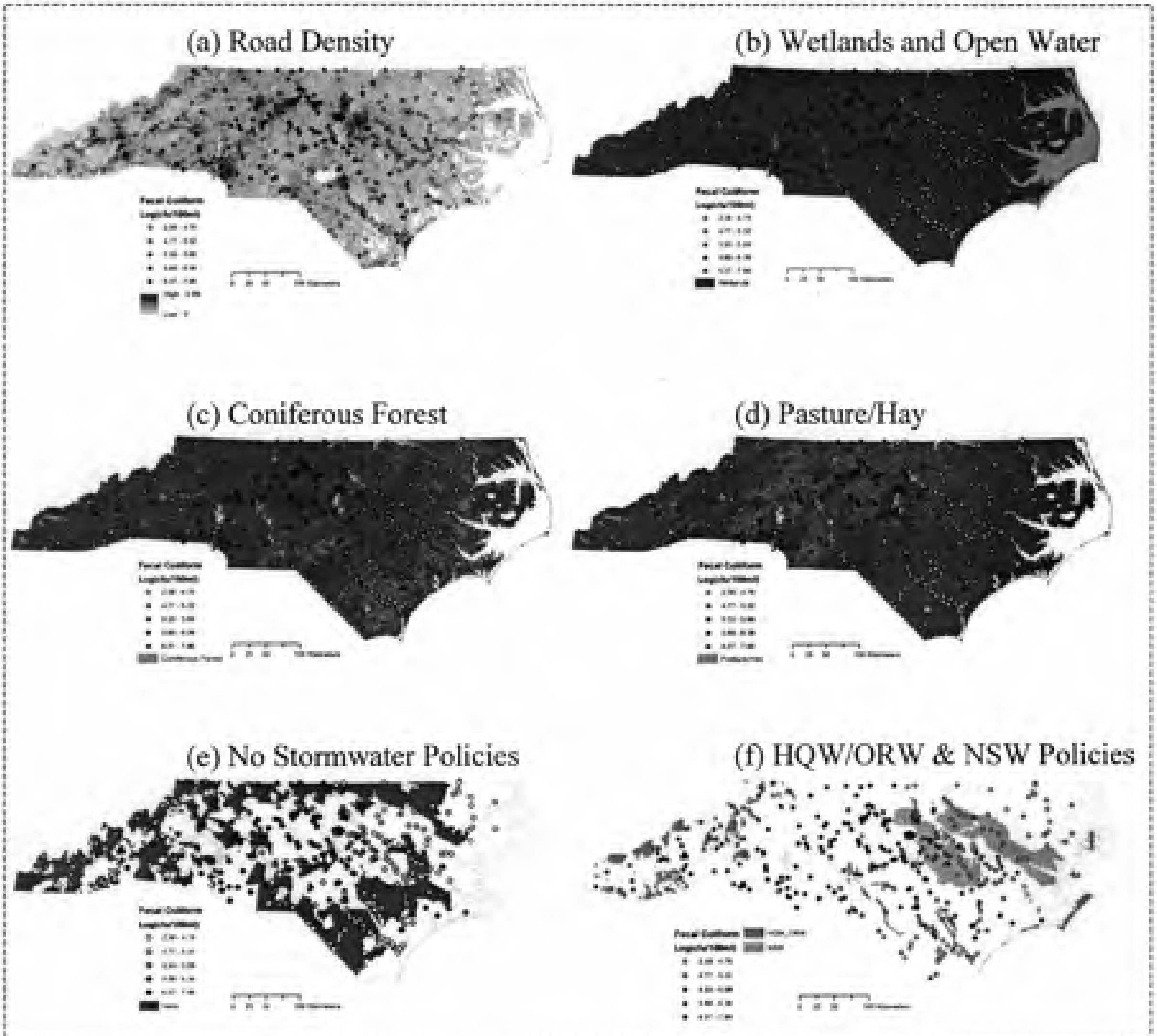
Figure 17. *Plots showing Analysis of the Residuals of the Multivariate Model*



5.4 Maps of Predictor Variables and Fecal Coliform

Figure 18 provides a series of maps which overlay log mean fecal coliform data at each monitoring location over the predictor variables which were selected in the multivariate regression model. These maps show the distinct regional patterns in both the fecal coliform data, as well as many of the explanatory variables.

Figure 18. *Maps of Mean Log-Fecal Coliform at Monitoring Locations and Overlap with Select Predictor Variables from the Multivariate Regression*



5.5 BMP Survey Results

Finally, an analysis of local post-construction BMP implementation practices was conducted for a 75 point subset of the 338 data points. The BMP Survey variables were regressed against the residuals derived from the predictive models. The only variable which was selected through this process was "charging fines for noncompliance in long-term BMP maintenance responsibilities". This variable was selected when using the best model from the forward, backward, and exhaustive methods. Forty-one of the 75 points (55%) were located within the boundaries of jurisdictions which charged fines for noncompliance at the time of the survey (2006), 31 points (41%) were within jurisdictions which did not charge fines, and three of the points had missing data for this variable. Of the 41 points where fines were charged, 30 only charged fines, while 11 charged fines in addition to another penalty. Charging fines for noncompliance was seen as a protective policy which was associated with lower fecal coliform levels. The coefficient for this variable was estimated at -0.462, and was significant ($P<0.01$). The variable accounted for about 11% of variation in the residuals.

The second method of analyzing the impact of BMP implementation practices was to rerun the regression with all independent variables for the 75 points that had BMP implementation data. However, this method did not result in any survey variables being selected.

6. Discussion

6.1 Land Use Regression and Statewide Stormwater Management Policies

The bivariate regression analysis indicates that all urban development variables are correlated with increased fecal coliform contamination, which is in accordance with previous studies. The swine lagoon density variable was shown to be associated with lower fecal coliform levels, which is counterintuitive, as it is a source of fecal matter. More analysis of this variable is needed to better understand the relationship with fecal coliform levels in North Carolina. One idea is to use the number or density of pigs instead.

Bivariate regressions of the individual land covers showed that the land covers which were correlated with lower fecal coliform levels were open water, natural barren land, shrub, cultivated crop cover, coniferous forest, and the two wetland classes. Wetlands had the largest R^2 of 0.30, indicating that about 30% of the variation in the fecal coliform levels was explained by this single variable. Open water had the next highest R^2 of 0.29. Perhaps these two variables have a high inverse relationship with fecal coliform due to dilution. After wetlands and open water, the next highest R^2 of the individual land classes was pasture/hay with 0.20.

Of the non-land cover variables, Phases I and II Policy had the highest R^2 of 0.22, followed by road density with an R^2 of 0.21. Both of these variables were correlated with increased levels of fecal coliform as they are indicators of increased urban development and impervious cover and higher, more polluted stormwater runoff. The two environmental variables analyzed were both inversely correlated with fecal coliform. The inverse relationship between temperature and fecal coliform is supported by the literature. As for precipitation, while individual storm events contribute to high loads of fecal coliform in streams, having relatively more precipitation on average could indicate areas with larger bodies of water, which can act to dilute fecal coliform concentrations.

Finally, a unique contribution of this study is the inclusion of the spatial coverage of federal and state stormwater management programs into the regression analysis. Results from the bivariate regressions indicate that stormwater management rules for Coastal, Nutrient Sensitive Waters, and High Quality and Outstanding Resource Waters are all correlated with decreases in fecal coliform levels. However, as mentioned previously, there is a distinct geographical pattern across North Carolina which indicates that fecal coliform decreases as it moves from the Piedmont region towards the coast. Thus, it is difficult to determine if this observed geographical pattern is due to land use, environmental conditions, or policies.

In fact, many stormwater management policy variables may only show correlations with fecal coliform due to their associations with urban development. For example, the NPDES Phase I and II policies, only apply to urbanized areas which are large enough to operate a MS4. Thus, these policies are highly correlated with the level of urban development, which has been shown to be associated with high concentrations of fecal coliform bacteria. In contrast, areas without any stormwater management regulations are associated with lower fecal coliform levels, since those are non-urbanized areas. Thus, the relationships between these policies and fecal coliform are due to urban development rather than any intrinsic characteristics of the policies. What is interesting is that when urban development is controlled for by road density in the multivariate model, certain stormwater policies were still selected as significant variables.

6.2 Multivariate Regression Model

Conducting a multivariate regression uses a combination of independent variables to explain fecal coliform and thus addresses the issue of leaving out potentially important confounding variables during the initial bivariate regressions. Since many of the variables in this study are correlated with each other, a variety of models are likely to have similar explaining power. In fact, the forward stepwise, backward stepwise, and exhaustive (or all possible subsets) methods each resulted in a slightly different model; however, five of the independent variables were selected by all three methods, indicating their robustness in explaining fecal coliform. In the future, cross-validation should also be employed to assess the robustness of the selected model in predicting fecal coliform concentrations in surface water. The final model reported was selected through the backward method, as that method produced the model with the fewest variables and with lowest multicollinearity without substantially sacrificing fit.

The results of the selected model agree with findings in previous literature that high levels of urban development, as measured by road density, lead to increased fecal coliform contamination. The model further agrees with past studies that agricultural land is correlated with higher fecal coliform contamination, due to high concentrations of grazing cattle and farm animals. The model also indicates that coniferous forested land cover is correlated with increased fecal coliform levels. This positive relationship was also seen in previous studies and was explained as being due to concentrations of wildlife in forested areas. Previous land use regression studies of fecal coliform contamination have not included the water variable in the model. Since location next to open water is a highly significant correlate in the model, as indicated by the second highest R^2 in the bivariate regression analysis, inclusion of this variable in the multivariate model could obscure the possible protective impacts of non-urban land covers.

Lastly, no other studies have incorporated stormwater management policies into a predictive model for surface water quality at the state level. The policies that were selected for the model were the Nutrient Sensitive Waters regulations and the High Quality and Outstanding Resource Waters Program. The High Quality and Outstanding Resource Waters policy variable is likely associated with lower fecal coliform contamination, since these waters have already been deemed as high quality by the state. The negative association between Nutrient Sensitive Waters and fecal coliform however poses interesting questions. More analysis is needed to identify why this relationship is seen. This finding is especially intriguing since Mallin et al. (2000) found that fecal coliform bacteria were positively correlated with nitrates in the water.

6.3 Local BMP Implementation Practices

As mentioned in the data section, the BMP Survey data from Barnes and Bruce (2008) may be biased and contain unrepresentative data for some jurisdictions which were in the early stages of developing a stormwater management program at the time of the survey. Thus, no definitive conclusions can be made from these results, but they are still interesting to contemplate.

The results from the two different analysis methods show an inverse relationship between fecal coliform levels and whether a jurisdiction charges fines for noncompliance with long-term BMP maintenance responsibilities. Perhaps this signifies that charging a fine on its own or in combination with another penalty can be a more effective strategy at incentivizing long-term BMP maintenance and thus improved water quality than having no penalties, only issuing stop work orders, or only employing other penalties.

While the results for local government practices of BMP implementation should be interpreted cautiously due to limitations in the data, they represent the type of questions that would be of interest to local and state planners concerned with stormwater management. If more post-construction BMP surveys are conducted and better data are collected in the future, more meaningful conclusions could be drawn. It is recommended that future studies undertake a more comprehensive survey of local government practices to manage stormwater runoff.

6.4 Limitations and Future Improvements

The model and methods represented in this technical report are not without limitations. As the maps of the land cover explanatory variables show, there is a distinct geographical pattern in the data. The model presented in this report does not have a variable which controls for region. Future state-wide studies should incorporate a variable to account for regional variations in fecal coliform. This variable could be continuous, such as distance from ocean, or a categorical variable such as coastal, piedmont, or mountain region.

The use of EPA STORET data allows for data analysis at a large geographic and temporal scale. However, a limitation to this data source is that it may not provide a random sample of locations. Many monitoring locations are designated in areas where water pollution is an existing problem or imminent threat. Thus, the fecal coliform data may be biased in representing more polluted surface waters.

Another limitation in the sampling methodology is the restriction of monitoring locations to include only those which were monitored at least nine times per year, since it excludes a lot of valuable data. Future studies could look at data for just the summer season and see if this increases the sample size. This is also important, since many recreational areas, the water quality of which is a public health concern, are those that are only monitored six times per year in the summer.

Another potential source of bias in the data could arise from different sampling procedures at different monitoring locations. While there are some basic guidelines which must be followed, there are still certain variables which depend on the discretion of the staff collecting the samples. For example, city staff members may decide the number of samples to be taken during storm events. Jurisdictions with more resources and knowledgeable staff are likely to follow better sampling protocol. This could be a confounding factor for fecal coliform, since areas with better reporting may also show higher levels of fecal coliform because of the that higher standard. It has been suggested that incorporating a variable on jurisdictional poverty or finances could help control for this.

A final limitation of this study is the use of circular buffers instead of nested watersheds. The benefits to the circular buffer approach are that the calculations are simpler and it allows for analysis at different spatial scales. Studies have also shown that the coefficients in the linear model do not significantly differ between the two approaches. However, a nested watershed approach is the common practice since it is hydrologically more accurate. Additionally, a watershed-based model is likely to have an improved R^2.

7. Conclusion

The ways that people use and manage land has been shown to be a primary cause of land-cover change, and thus serve as a fundamental source of change in surface water quality. Land use planners have the power to control the extent and location of land-cover changes by influencing the pattern of urban development in their jurisdictions. Furthermore, planning documents and ordinances can require mitigation programs (such as post-construction BMPs) to supplement land use patterns.

This research has provided further evidence that different land use characteristics, level of impervious land cover, and stormwater management programs have an impact on water quality across North Carolina. The findings that road density, percentage impervious cover, and grazing land cover lead to increased levels of fecal coliform are in accordance with other previous studies. Results indicate that wetlands and open water are associated with decreased fecal coliform levels, likely due to a dilution effect. However, lower fecal coliforms near the coast, where there are more open bodies of water, could also be indicative of the known fact that increased salinity decreases fecal coliform levels or because there is less urban development in the coastal region.

This research is the first to incorporate state and federal policies within a predictive model for surface water quality. The findings indicate that NPDES Phase I and Phase II regulations predict increased levels of fecal coliform, since they are directly related with population density. On the other hand, proximity to lands which fall under the high quality and outstanding resource water stormwater regulations and nutrient sensitive water regulations are associated with lower fecal coliform levels. Furthermore, including these stormwater policies improved model R^2 by 0.05 or about 5%.

This report represents a first step in analyzing the effectiveness of different local implementations practices of stormwater policies across different jurisdictions in the state. The conclusions that can be drawn from these initial findings are limited by the little available data on post-construction BMP implementation for different jurisdictions. Results point towards associations between lower fecal coliform levels and charging fees for noncompliant BMPs as well as higher fecal coliform levels in locations that do not keep a centralized inventory of BMPs. However, much more research is needed to produce high quality and reliable results that can be used to guide planners in North Carolina on the types of land use and stormwater management practices that will work best to mitigate impacts of NPS fecal contaminant runoff to surface water in the future.

Additional future research focuses may include: (1) using alternative measures of explanatory variables, especially the swine lagoon variable, and including new explanatory variables, such as information on jurisdictional finances and a variable to control for regional variations in fecal coliform; (2) testing if this land-use regression model improves fecal coliform estimation at unmonitored locations over simpler interpolation techniques; and (3) recreating the model using nested watersheds instead of circular buffers to improve model fit.

References

Arnold, C. L. and Gibbons,C. J. (1996). Impervious surface coverage: The emergence of a key environmental indicator. *Journal of the American Planning Association, 62*(2), 243-258.

Aufdenakmpe, A. K., Arscott, D. B., Dow, C. L., and Standley, L. J. (2006). Molecular Tracers of Soot and Sewage Contamination in Streams Supplying New York City Drinking Water. *Journal of the North American Benthological Society, 25*(4), 928-953.

Baker, A. (2005). 188: Land use and water quality. In M. G. Anderson (Ed.), *Encyclopedia of Hydrological Sciences* (1-6).

Barnes, G. and Bruce. S. (2008). Survey of Local Government Post-Construction BMP Maintenance and Enforcement in North Carolina: Report of Findings.

Brabec, E., Schulte, S., Richards, P. L. (2002). Impervious surfaces and water quality: A review of current literature and its implications for watershed planning. *Journal of Planning Literature, 16*(4), 499-514.

Cahoon, L. B., Hales, J. C., Carey, E. S., Loucaides, S., Rowland, K. R., Nearhoof, J. E. (2006). Shellfishing closures in southwest Brunswick County, North Carolina: Septic tanks vs. storm-water runoff as fecal coliform sources. *Journal of Coastal Research, 22*(2), 319-327.

Craney, T. A. and Surles, J. G. (2002). Model-Dependent Variance Inflation Factor Cutoff Values. *Quality Engineering, 14*(3), 391-403.

Eleria, A. and Vogel, R. M. (2005). Predicting fecal coliform bacteria levels in the Charles River, Massachusetts, USA. *Journal of the American Water Resources Association, 41*(5), 1195-1209.

EPA. (2009). National Water Quality Inventory: Report to Congress 2004 Reporting Cycle Retrieved from: http://water.epa.gov/lawsregs/guidance/cwa/305b/upload/2009_01_22_305b_2004report_2004_305Breport.pdf

EPA. (2010). The National Water Quality Assessment Report for North Carolina. Retrieved from: http://iaspub.epa.gov/waters10/attains_state.control?p_state=NC

EPA. (2012). Water Monitoring and Assessment: 5.11 Fecal Bacteria. Retrieved from: http://water.epa.gov/type/rsl/monitoring/vms511.cfm

Fisher, D. S., Steiner, J. L., Endale, D. M., Stuedemann, J. A., Schomberg, H. H., Franzluebbers, A. J., Wilkinson, S. R. (2000). The relationship of land use practices to surface water quality in the Upper Oconee Watershed of Georgia. *Forest Ecology and Management, 128*, 39–48.

Frenzel, S. A. and Couvillion, C. S. (2002). Fecal-indicator bacteria in streams along a gradient of residential development. *Journal of the American Water Resources Association, 38*(1), 265-273.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., (2011). Completion of the 2006 National Land Cover Database for the Conterminous United States, *PE&RS*, Vol. 77(9):858-864.

Goonetilleke, A., Thomas, E., Ginn, S., Gilbert, D. (2005). Understanding the role of land use in urban stormwater quality management. *Journal of Environmental Management 74*, 31–42.

Harclerode, C. L., Gentry, T. J., Aitkenhead-Peterson, J. A. (2012). A geographical approach to tracking Escherichia coli and other water quality constituents in a Texas coastal plains watershed. Environmental Monitoring and Assessment.

Kang, J., Lee, S. W., Cho, K. H., Ki, S. J., Cha, S. M., Kim, J. H. (2010). Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin. *Water Research, 44*, 4143-4157.

Kelsey, H., Porter, D. E., Scott, G., Neet, M., White, D. (2004). Using geographic information systems and regression analysis to evaluate relationships between land use and fecal coliform bacterial pollution. *Journal of Experimental Marine Biology and Ecology, 298*, 197– 209.

LeSage, J. P. (2004). Lecture 1: Maximum likelihood estimation of spatial regression models. Retrieved form: http://www4.fe.uc.pt/spatial/doc/lecture1.pdf

Line, D. E. (2002). Changes in land use/management and water quality in the Long Creek Watershed. *Journal of the American Water Resources Association, 38*(6), 1691-1701.

Lohse, K. A. and Merenlender, A. M. (2009). Impacts of exurban development on water quality. In A.X. Esparza, G. McPherson (eds.), *The Planner's Guide to Natural Resource Conservation* (159-179).

Mallin, M. A., Williams, K. E., Esham, E. C., Lowe, R. P. (2000). Effect of human development on bacteriological water quality in coastal watersheds. *Ecological Applications, 10*(4), 1047–1056.

Mallin, M. A., Ensign S. H., McIver, M. R., Shank, G. C., Fowler, P. K. (2001). Demographic, landscape, and meteorological factors controlling the microbial pollution of coastal waters. *Hydrobiologia, 460*, 185–193.

Mallin, M. A., Johnson, V. L., Ensign, S. H. (2009). Comparative impacts of stormwater runoff on water quality of an urban, a suburban, and a rural stream. *Environmental Monitoring and Assessment, 159*, 475–491.

North Carolina Division of Water Quality (NCDWQ) (2009). *Stormwater Best Management Practices Manual.* Retrieved from: http://portal.ncdenr.org/web/wq/ws/su/bmp-manual

Peierls, B. L., Caraco, N. F., Pace, M. L., Cole, J. J. (1991). Human influence on river nitrogen. *Nature, 350*(6317), 386-387.

Petersen, T. M., Rifai, H. S., Suarez, M. P., Stein, A. R. (2005). Bacteria loads from point and nonpoint sources in an urban watershed. *Journal of Environmental Engineering, 131*(10), 1414-1425.

Pratt, B. and Chang, H. (2012). Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials 209-210*, 48-58.

Pruss, A. (1998). Review of epidemiological studies on health effects from exposure to recreational water. *International Journal of Epidemiology, 27*, 1-9.

Ren, W., Zhong, Y., Meligrana, J., Anderson, B., Watt, W. E., Chen, J., Leung, H. (2003). Urbanization, land use, and water quality in Shanghai 1947–1996. *Environment International 29*, 649– 659.

Santhi, C., Srinivasan, R., Arnold, J. G., Williams, J. R. (2006). A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. *Environmental Modelling & Software, 21*, 1141-1157.

Schueler, T. R., Fraley-McNeal, L., Cappiella, K. (2009). Is impervious cover still important? Review of recent research. *Journal of Hydrologic Engineering, 14*(4), 309-315.

Selvakumar, A. and Borst, M. (2006). Variation of microorganism concentrations in urban stormwater runoff with land use and seasons. *Journal of Water and Health, 4*(1), 109-124.

Sliva, L. and Williams, D. D. (2001). Buffer zone versus whole catchment approaches to studying land use impact on river water quality. *Water Research, 35*(14), 3462-3472.

Smith, J. H., Wickham, J. D., Norton, D., Wade, T. G., Jones, K. B. (2001). Utilization of landscape indicators to model potential pathogen impaired water. *Journal of the American Water Resources Association, 37*(4), 805-814.

Tong, S. T. Y., and Chen, W. (2002). Modeling the relationship between land use and surface water quality. *Journal of Environmental Management, 66*, 377-393.

Tu, J. (2011). Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. *Applied Geography, 31*, 376-392.

U.S. Department of Agriculture (USDA). (2007). The 2007 Census of Agriculture: Hogs and Pigs Fact Sheet. Retrieved from: http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Fact_Sheets/Production/hogsandpigs.pdf

Wang, X. (2001). Integrating water-quality management and land-use planning in a watershed context. *Journal of Environmental Management, 61*, 25–36.

Yang, X. and Jin, W. (2010). GIS-based spatial regression and prediction of water quality in river networks: A case study in Iowa. *Journal of Environmental Management, 91*, 1943-1951.

Young, K. D. and Thackston, E. L. (1999). Housing density and bacterial loading in urban streams. *Journal of Environmental Engineering, 125*(12), 1177-1180.